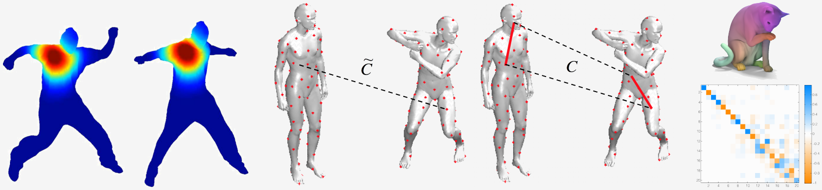
Sparse Modeling of Intrinsic Correspondences

Recent Advances in the Analysis of 3D Shapes

Tobias Gurdan
Dr. Emanuele Rodolà

June 18, 2014



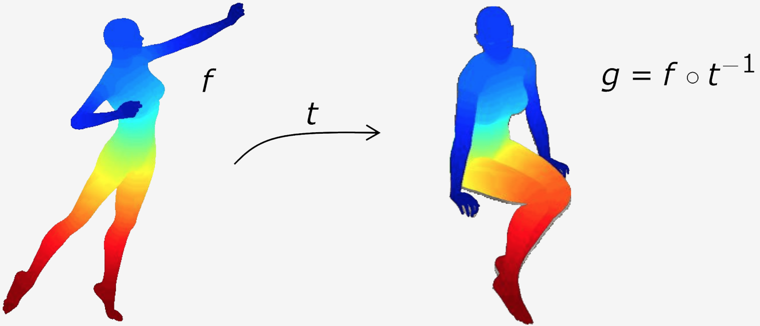
(a) Kernels (b) Assignment Problem (c) FM

Figure : Topics from shape analysis we have already seen

Figure : Two sets of regions with unknown correspondence as input for sparse modeling. The method we're going to see now.

Introduction

- Given two shapes X and Y (compact smooth Riemannian manifolds)
- Let $t : X \to Y$ be a bijection between them
- Then, for any real function $f : X \to \mathbb{R}$, we can construct a corresponding function $g : Y \to \mathbb{R}$ as $g = f \circ t^{-1}$



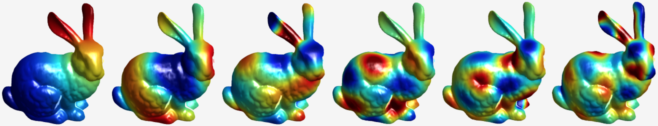
The Detour

- The correspondence t uniquely defines a mapping between two function spaces
- We introduce $T(f) = f \circ t^{-1}$ which maps functions in X to functions in Y
- More formally $T : \mathcal{F}(X, \mathbb{R}) \to \mathcal{F}(Y, \mathbb{R})$, where $\mathcal{F}(X, \mathbb{R})$ denotes the space of real functions on X
- Such a representation is linear, since for every pair of functions f_1, f_2 and scalars α_1, α_2,

$$\begin{aligned} T(\alpha_1 f_1 + \alpha_2 f_2) &= (\alpha_1 f_1 + \alpha_2 f_2) \circ t^{-1} \\ &= \alpha_1 f_1 \circ t^{-1} + \alpha_2 f_2 \circ t^{-1} \\ &= \alpha_1 T(f_1) + \alpha_2 T(f_2). \end{aligned} \tag{1}$$

Basis Representations

- Assume that X and Y are equipped with bases $\{\phi_i\}_{i\geq 1}$ and $\{\psi_j\}_{j\geq 1}$ respectively



- Any $f : X \to \mathbb{R}$ can be represented as

$$f = \sum_{i\geq 1} a_i \phi_i \tag{2}$$

- Likewise, for any $g : Y \to \mathbb{R}$

$$g = \sum_{j\geq 1} b_j \psi_j \tag{3}$$

Basis Transformations

- Let's see how we can express g in terms of our bases on X.

$$g = T(f) = T\left(\sum_{i\geq 1} a_i \phi_i\right) = \sum_{i\geq 1} a_i T(\phi_i) \tag{4}$$

- Also, for every i there exists coefficients c_{ij} such that

$$T(\phi_i) = \sum_{j\geq 1} c_{ij}\psi_j, \tag{5}$$

- Plug 5 in 4.

$$T(f) = \sum_{i,j\geq 1} a_i c_{ij}\psi_j. \tag{6}$$

Discretize ALL the Things

- Let us now assume finite sampling of X and Y with m samples
- The bases are represented as the $m \times n$ matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ containing, respectively, n discretized functions ϕ_i and ψ_j as the columns
- Thus for f and $g = T(f)$ with coefficients $\mathbf{a}$ and $\mathbf{b}$

$$f = \sum_{i \geq 1} a_i \phi_i \rightarrow \mathbf{f} = \mathbf{\Phi} \mathbf{a}$$

$$g = \sum_{j \geq 1} b_j \psi_j \rightarrow \mathbf{g} = \mathbf{\Psi} \mathbf{b}$$

Amazingly Efficient Coding

- Using this notation, Equation (6) can be rewritten as

$$\mathbf{g} = \mathbf{\Psi b} = T(\mathbf{f}) = T(\mathbf{\Phi a}) = \mathbf{\Psi C}^{\mathrm{T}}\mathbf{a}; \tag{7}$$

- Since $\mathbf{\Psi}$ is invertible, this simply means that

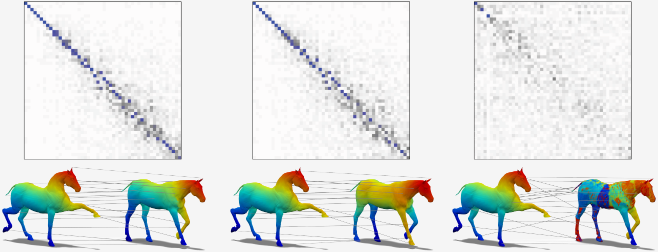
Functional Map

$$\mathbf{b}^{\mathrm{T}} = \mathbf{a}^{\mathrm{T}}\mathbf{C} \tag{8}$$

- Thus, the $n \times n$ matrix $\mathbf{C}$ fully encodes the linear map T between the functional spaces
- It contains the coordinates in the basis $\mathbf{\Psi}$ of the mapped elements of the basis $\mathbf{\Phi}$

Point-to-Point Correspondence

- Each point on X correpsponds to some point on Y
- In functional representation: Each row of $\mathbf{\Psi}\mathbf{C}^{\mathrm{T}}$ coincides with some row of $\mathbf{\Phi}$
- Solvable using Iterative Closest Points (ICP) in n dimensions
- We can simply recover T by means of $\mathbf{P} = \mathbf{\Phi}\mathbf{C}^{T}\mathbf{\Psi}^{-1}$
- The discretized harmonic bases $\mathbf{\Phi}$ and $\mathbf{\Psi}$ yield nearly-diagonal and therefore very sparse C



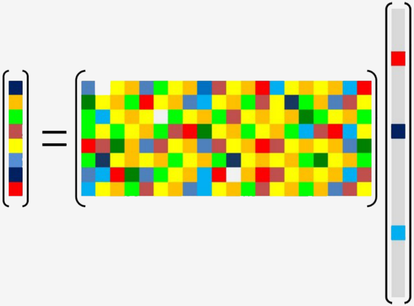
- Let's use this to our advantage!

Basics

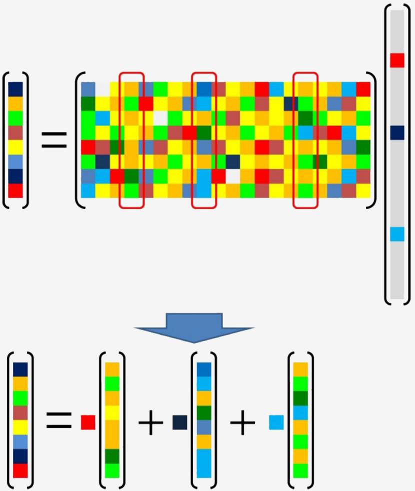
- Basic idea: represent a signal $\mathbf{x}$ using a compact (i.e. sparse) linear combination with coefficients $\mathbf{z}$ in an appropriate domain $\mathbf{D}$

$$\mathbf{x} = \mathbf{D}\mathbf{z} \tag{9}$$

- $\mathbf{D}$ is called the *dictionary*
- It is often selected to be *overcomplete*



Basics



- Toy example: Languages with alphabets, words and sentences

Dictionary Learning

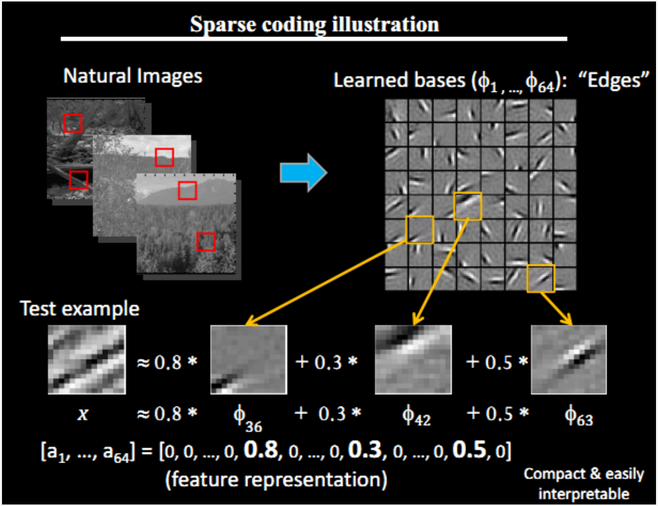


Figure : Sparse coding in image processing

Dictionary Learning

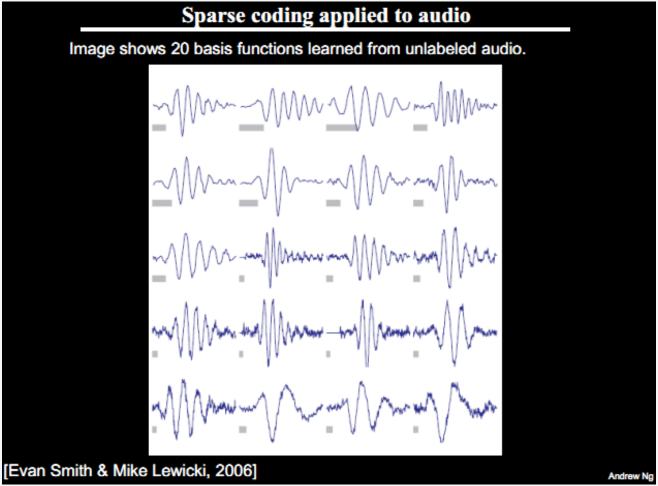


Figure : Sparse coding in audio processing

Dictionary Learning

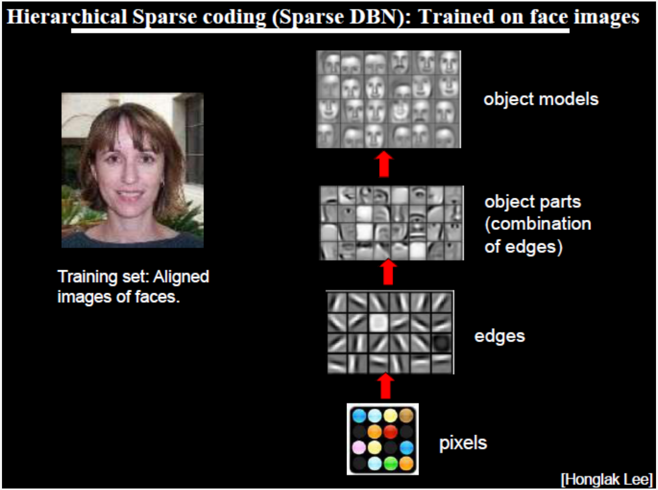


Figure : Example of dictionary learning on face images

Problem Formulation

- How can we find a compact (i.e. sparse) representation $\mathbf{z}$ of a signal $\mathbf{x}$ in a given dictionary $\mathbf{D}$? (Q)
- This is known as sparse representation *pursuit* or *sparse coding* and can be solved using the so-called Lasso formulation.

$$\min_{\mathbf{z}} \|\mathbf{x} - \mathbf{D}\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_1. \qquad (10)$$

Figure : p-norm for $p \in 2, 1, 0.5, 0.1$ from left to right.

Problem Formulation

- It is often useful to consider the structure shared by multiple vectors
- This is called the *collaborative* sparse model where we optimize

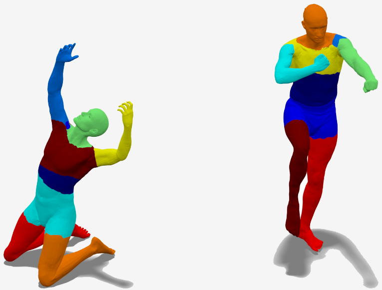
$$\min_{\mathbf{Z}} \|\mathbf{X} - \mathbf{D}\mathbf{Z}\|_{\mathrm{F}}^2 + \lambda \|\mathbf{Z}\|_{2,1}, \tag{11}$$

- The $\ell_{2,1}$ norm is defined as $\|\mathbf{Z}\|_{2,1} = \|\mathbf{z}_1^{\mathrm{T}}\|_2 + \cdots + \|\mathbf{z}_m^{\mathrm{T}}\|_2$, where $\mathbf{z}_i^{\mathrm{T}}$ denotes the *i*-th row of $\mathbf{Z}$ and promotes row-wise sparsity of the solution

Figure : Image reconstruction using collaborative sparse coding.

From FM's to PSC

- Let's assume, we have some region or feature detection process, i.e. maximally stable components (MSER)
- Given X it produces a collection of q functions $f_i : X \to \mathbb{R}$ based on the intrinsic properties of the shape only
- For now, we also assume that it is perfectly repeatable
- Thus it should also produce a collection of q similar functions $g_j : Y \to \mathbb{R}$



From FM's to PSC

- For every f_i there exists $g_j = f_i \circ t$ for unknown t
- Note, that we do not know the ordering, i.e. which f_i corresponds to which g_j
- Again $\mathbf{f}_i = \mathbf{\Phi a}_i$ and $\mathbf{g}_j = \mathbf{\Psi b}_j$
- $\mathbf{b}_j^{\mathrm{T}} = \mathbf{a}_i^{\mathrm{T}}\mathbf{C}$ (8) must still hold for all i and j
- Row-wise stacking the coefficient vectors $\mathbf{a}_i$ and $\mathbf{b}_j$ to $q \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ we can write

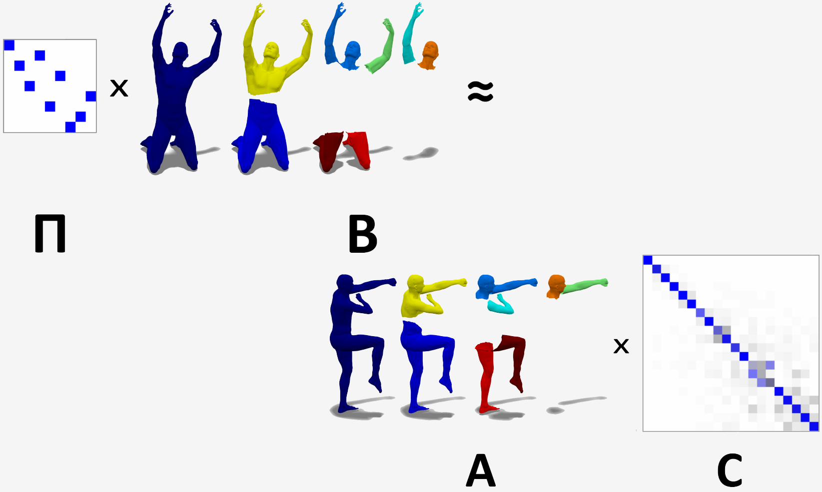
$$\mathbf{B} = \mathbf{AC}, \tag{12}$$

- We want diagonality of $\mathbf{C}$, so we introduce a $q \times q$ permutation matrix $\mathbf{\Pi}$.

$$\mathbf{\Pi B} = \mathbf{AC}, \tag{13}$$

with $\pi_{ij} = 1$ if $\mathbf{a}_i$ corresponds to $\mathbf{b}_j$

From FM's to PSC



From FM's to PSC

- Recall the sparse representation pursuit where we looked for $\mathbf{X} = \mathbf{D}\mathbf{Z}$.

$$\min_{\mathbf{Z}} \|\mathbf{X} - \mathbf{D}\mathbf{Z}\|_{\mathrm{F}}^2 + \lambda \|\mathbf{Z}\|_1$$

- Now let's look at our problem.

Permuted Sparse Coding

$$\min_{\mathbf{C},\mathbf{\Pi}} \frac{1}{2} \|\mathbf{\Pi}\mathbf{B} - \mathbf{A}\mathbf{C}\|_{\mathrm{F}}^2 + \lambda \|\mathbf{W} \odot \mathbf{C}\|_1 \tag{14}$$

- $\mathbf{W}$ has small weights on diagonal and large weights on off-diagonal elements, promoting diagonal solutions of $\mathbf{C}$

From FM's to PSC

- How to solve this?
- Let's fix $\mathbf{\Pi}$ and set $\mathbf{B}' = \mathbf{\Pi}\mathbf{B}$.

$$\min_{\mathbf{C}} \frac{1}{2}\|\mathbf{B}' - \mathbf{A}\mathbf{C}\|_{\mathrm{F}}^2 + \lambda\|\mathbf{W} \odot \mathbf{C}\|_1, \tag{15}$$

- This resembles the Lasso problem seen in the sparse representation pursuit!
- Now let's fix $\mathbf{C}$ and set $\mathbf{A}' = \mathbf{A}\mathbf{C}$.

$$\min_{\mathbf{\Pi}} \frac{1}{2}\|\mathbf{\Pi}\mathbf{B} - \mathbf{A}'\|_{\mathrm{F}}^2 + \lambda\|\mathbf{W} \odot \mathbf{C}\|_1 \tag{16}$$

$$\equiv \min_{\mathbf{\Pi}} \frac{1}{2}\|\mathbf{\Pi}\mathbf{B} - \mathbf{A}'\|_{\mathrm{F}}^2 \tag{17}$$

- This is exactly the linear assignment problem!

From FM's to PSC

- We can rewrite

$$\|\mathbf{\Pi B} - \mathbf{A}'\|_{\mathrm{F}}^2 = \mathrm{tr}\left(\mathbf{B}^{\mathrm{T}}\mathbf{\Pi}^{\mathrm{T}}\mathbf{\Pi B}\right) - 2\mathrm{tr}\left(\mathbf{B}^{\mathrm{T}}\mathbf{\Pi}^{\mathrm{T}}\mathbf{A}'\right) + \mathrm{tr}\left(\mathbf{A}'^{\mathrm{T}}\mathbf{A}'\right). \tag{18}$$

- Since $\mathbf{\Pi}^{\mathrm{T}}\mathbf{\Pi} = \mathbf{I}$ we get

$$\max_{\mathbf{\Pi}} \mathrm{tr}\left(\mathbf{\Pi}^{\mathrm{T}}\mathbf{E}\right), \tag{19}$$

  where $\mathbf{E} = \mathbf{A}'\mathbf{B}^{\mathrm{T}} = \mathbf{ACB}^{\mathrm{T}}$

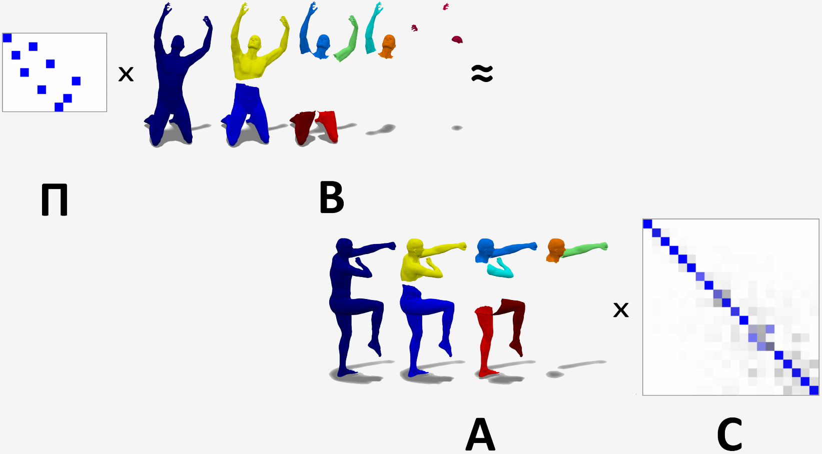
- To make it practically solvable, we allow $\mathbf{\Pi}$ to be doubly stochastic.

$$\max_{\mathbf{\Pi} \geq \mathbf{0}} \ \mathrm{vec}(\mathbf{E})^{\mathrm{T}}\mathrm{vec}(\mathbf{\Pi})\mathrm{s.t.} \quad \begin{cases} \mathbf{\Pi 1} = \mathbf{1} \\ \mathbf{\Pi}^{\mathrm{T}}\mathbf{1} = \mathbf{1}. \end{cases} \tag{20}$$

Robust Permuted Sparse Coding

- What if we don't have a bijective correspondence betweeen the functions?
- Assume q functions f_i detected on X and r functions g_j detected on Y $(q \leq r)$
- We now have a $q \times r$ partial permutation matrix $\mathbf{\Pi}$
- We can model the non-surjectivity of $\mathbf{\Pi}$ by relaxing $\mathbf{\Pi}^{\mathrm{T}}\mathbf{1} = \mathbf{1}$ to $\mathbf{\Pi}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}$

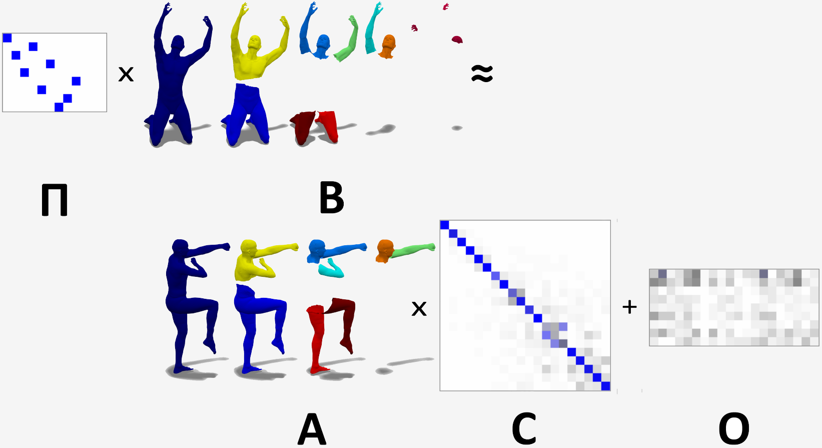
Robust Permuted Sparse Coding



Robust Permuted Sparse Coding

- Generally, only $s \leq q$ f_i's have corresponding g_j's
- Only an unknown subset of rows of $\mathbf{\Pi B} \approx \mathbf{AC}$ hold
- Drop injectivity too and relax $\mathbf{\Pi 1} = \mathbf{1}$ to $\mathbf{\Pi 1} \leq \mathbf{1}$
- But wait! (Q) Trivial solution $\mathbf{\Pi} = \mathbf{0}$
- Idea: absorb the $r - s$ mismatches in a row-sparse $q \times n$ outlier matrix $\mathbf{O}$

Robust Permuted Sparse Coding



Robust Permuted Sparse Coding

- The $q \times r$ matrix $\mathbf{\Pi}$ is searched over all injective correspondences

Robust Permuted Sparse Coding

$$\min_{\mathbf{C},\mathbf{O},\mathbf{\Pi}} \frac{1}{2}\|\mathbf{\Pi}\mathbf{B} - \mathbf{A}\mathbf{C} - \mathbf{O}\|_{\mathrm{F}}^2 + \lambda\|\mathbf{W} \odot \mathbf{C}\|_1 + \mu\|\mathbf{O}\|_{2,1} \tag{21}$$

- Again, we can split it into two sub-problems
- One with the fixed permutation $\mathbf{\Pi}$ where $\mathbf{B}' = \mathbf{\Pi}\mathbf{B}$.

$$\min_{\mathbf{C},\mathbf{O}} \frac{1}{2}\|\mathbf{B}' - \mathbf{A}\mathbf{C} - \mathbf{O}\|_{\mathrm{F}}^2 + \lambda\|\mathbf{W} \odot \mathbf{C}\|_1 + \mu\|\mathbf{O}\|_{2,1}, \tag{22}$$

- The other one with the fixed $\mathbf{C}$ where $\mathbf{E} = (\mathbf{A}\mathbf{C})\mathbf{B}^{\mathrm{T}}$.

$$\max_{\mathbf{\Pi} \geq \mathbf{0}} \ \mathrm{vec}(\mathbf{E})^{\mathrm{T}}\mathrm{vec}(\mathbf{\Pi})\mathrm{s.t.} \quad \begin{cases} \mathbf{\Pi}\mathbf{1} = \mathbf{1} \\ \mathbf{\Pi}^{\mathrm{T}}\mathbf{1} \leq \mathbf{1}, \end{cases} \tag{23}$$

Numerical Solution

- All sub-problems can be efficiently solved using established optimization algorithms
- Namely, 20 & 23 and 15 & 22 can be solved using the Hungarian algorithm and forward-backward splitting algorithms respectively.
- Pok et al. adapt the latter to the concrete setting. For convenience, here is their algorithm.

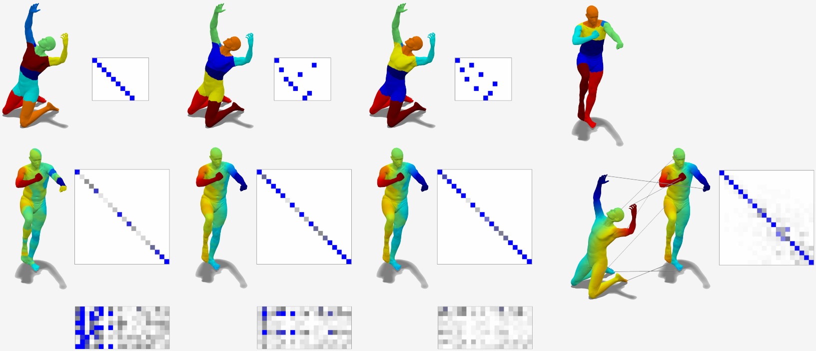
input : Data $\mathbf{B}', \mathbf{A}$; parameters λ, μ; step size α.
output : Sparse matrix $\mathbf{O}$ and row-wise sparse outlier matrix $\mathbf{O}$
Initialize $\mathbf{O}^0 = \mathbf{B}'$ and $\mathbf{C}^0 = \mathbf{0}$.
for *k=1,2,. . . ,until convergence* **do**

$$\mathbf{C}^{k+1} = \mathbf{P}_1\left((\mathbf{I} - \tfrac{1}{\alpha}\mathbf{A}^{\mathrm{T}}\mathbf{A})\mathbf{C}^k - \tfrac{1}{\alpha}\mathbf{A}^{\mathrm{T}}(\mathbf{O}^k - \mathbf{B}')\right)$$
$$\mathbf{O}^{k+1} = \mathbf{P}_2\left((1 - \tfrac{1}{\alpha})\mathbf{O}^k - \tfrac{1}{\alpha}(\mathbf{A}\mathbf{C}^k - \mathbf{B}')\right)$$

end

Convergence

- Fast and robust
- Here are three iterations (from left to right), showing $\mathbf{\Pi}_i$, $\mathbf{C}_i$ and $\mathbf{O}_i$ from top to bottom with according ordering and correspondences.

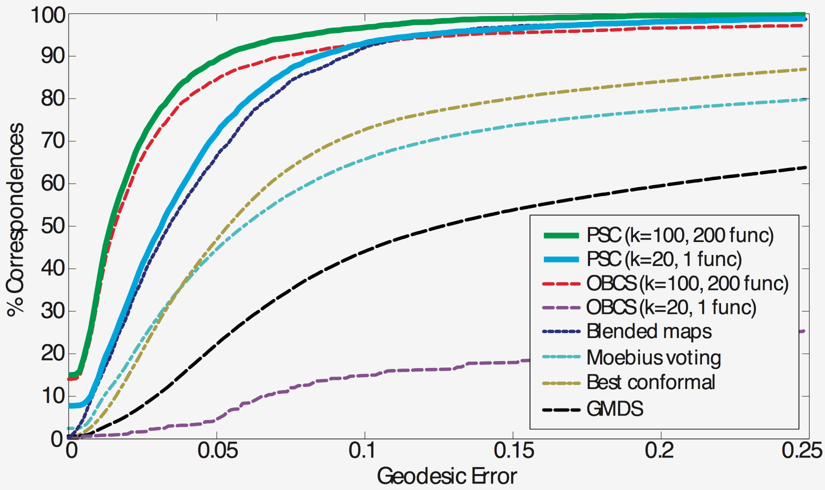


Average Runtime

- For the following results they used 20 eigenfunctions
- 5-15 detected regions with at least 5-10% of total shape area
- Initialize $\mathbf{C} = \mathbf{0}$ and $\mathbf{\Pi} = \frac{1}{q}\mathbf{1}\mathbf{1}^{\mathrm{T}}$

Vertices	Basis	MSER	Opt.	Ref.	Tot.
5K	0.53	0.61	7.80	1.41	**10.35**
10K	0.99	1.32	7.91	2.70	**12.92**
20K	2.03	3.58	7.91	5.52	**19.04**
50K	5.57	14.23	7.85	13.99	**41.64**

Comparison



100
90
80
70
60
50
40
30
20
10
0
% Correspondences
0
0.05
0.1
0.15
0.2
0.25
Geodesic Error
PSC (k=100, 200 func)
PSC (k=20, 1 func)
OBCS (k=100, 200 func)
OBCS (k=20, 1 func)
Blended maps
Moebius voting
Best conformal
GMDS

Pictuuures

Figure : Near-isometric deformations of the TOSCA male

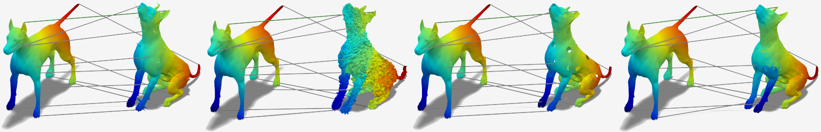


Figure : SHREC dog undergoing nearly-isometric deformations (spike noise, Gaussian noise, topological noise)

Limitations

Figure : The approach fails for significantly non-isometric shapes (Q)

The End

Thanks! :)